Abstract

Recent years have seen a rapid development of multimedia communication technologies which facilitate our life, but at the same time put security of digital audio at many risks, such as copyright infringement, malicious tampering. Audio data hiding techniques, in which special codes are embedded into actual audio content without any affection to its normal use, have been proposed as a potential solution for these issues. In general, audio data hiding methods must satisfy five requirements: (i) inaudibility—keeping watermark imperceptible to users, (ii) blindness—avoiding using double storage and communication channels, (iii) robustness—preventing intentional attacks from illegal users, (iv) high capacity—conveying a large amount of data, and (v) high reliability—precisely detecting the watermark.

Although most reported methods could partly satisfy the requirements, the trade-off among the requirements is still challenging. To keep the watermark inaudible, it is straightforward that perceptually insensitive features of audio signals should be exploited for embedding. The resistance against modifications such as lossy compression becomes weak since the hidden data could be easily destroyed without degrading the sound quality. Finding out suitable acoustic features to ensure both inaudibility and robustness simultaneously is one of the most important tasks of audio data hiding design.

The aim of this research is to propose an audio data hiding method that achieves a reasonable trade-off among the requirements and is applicable in practical problems. This research introduces a concept of dynamic phase manipulation for audio watermarking in which the human sound-perception mechanism and sophisticated embedding rule are utilized to solve the conflict among the requirements. First, the dynamic phase manipulation scheme finds out frequency components that are insensitive to human ears and resistant against signal processing operations to keep hidden data inaudible and robust. Second, an appropriate embedding rule is employed to account for blindness and high embedding capacity. Accordingly, the proposed method of audio data hiding could obtain the inaudibility and the robustness simultaneously. The proposed method is then applied to three typical applications: copy prevention, annotated audio, and information carrier over AM radio broadcast.

The phase manipulation technique is used to embed a bit of data into an audio signal by changing the phase according to two phase patterns. The amount of phase modification is an important factor which directly decides the performance of the data hiding system on the inaudibility and the robustness. The smaller amount keeps the hidden data less audible but more weak against processing and vice versa. The main goal of this research is to find out a region of frequency components suitable for embedding and corresponding amount of phase-modification for the embedding region based on the characteristics of human auditory system (HAS) and the variability of audio signals.

According to these considerations, the dynamic phase manipulation scheme is constructed as follows. Original audio is firstly analyzed to find out suitable frequency region for embedding. The phase modification of a frequency component cause distortion in a manner that is directly proportional to the magnitude of that component. Therefore, the amount of phase-modification should also be adapted to the magnitude. The amount of

phase-modification is determined based on the energy of the embedding region. The modified phase spectrum and original magnitude spectrum are processed to yield a watermarked signal. In data extraction process, the same analysis steps are performed to identify the embedding frequency components and the amount of phase-modification. Watermark decoder is then performed on the phase of these components to extract embedded data. Experimental results have shown that variant amount of phase-modification improves performance on inaudibility and robustness remarkably compared with the case that a fixed amount of phase-modification is used. It suggested that the dynamic phase manipulation scheme is effective for audio data hiding.

The proposed framework ensures the inaudibility and the robustness by exploiting the advantages human perception mechanism and the variability of audio signals. The blindness and the embedding capacity are achieved by the nature of the embedding rules. To combat against cropping and shifting attacks, the proposed framework is built with a frame synchronization scheme. The frame synchronization is performed by searching the starting point around the size of one frame. A correct starting point is detected when the confidence of extracting a bit is the highest among all the points in that frame. Confidentiality is ensured by incorporating security parameters. Watermark is encrypted with a secret key before being embedded into audio signal.

The proposed framework is evaluated with respect to inaudibility, robustness, blindness, and embedding capacity. The inaudibility is confirmed by objective difference grade and subjective listening test. The robustness is confirmed by the accuracy of extracted data against signal processing operations and attacks. Subjective test and robustness test are also carried out to confirm the effectiveness of the dynamic phase manipulation scheme. Bit rate is varied to investigate embedding capacity as well. The proposed audio data hiding method is then applied to protecting digital audio, audio entertainment, and information carrier over AM radio broadcast.

Keywords: audio data hiding, robustness, reliability, adaptive phase modulation, quantization index modulation.